

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- FOI b6 b7C b7D
- (i) APPLICANT: Wang, Elizabeth A.  
Rosen, Vicki A.  
Wozney, John M.
- (ii) TITLE OF INVENTION: Novel BMP Products
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.  
(B) STREET: 87 CAMBRIDGE PARK DRIVE  
(C) CITY: CAMBRIDGE  
(D) STATE: MA  
(E) COUNTRY: USA  
(F) ZIP: 02140
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: US  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: KAPINOS, ELLEN J.  
(B) REGISTRATION NUMBER: 32,245  
(C) REFERENCE/DOCKET NUMBER: GI5160C
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 617-876-1170  
(B) TELEFAX: 617-876-5851

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 592 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: BOS TAURUS
- (vii) IMMEDIATE SOURCE:

(A) LIBRARY: ~~XXXXXXXXXX~~ VINE GENOMIC IN LAMBDA J1  
(B) CLONE: LAMBDA BP-21

) POSITION IN GENOME:

(C) UNITS: bp

x) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..390

ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION: 46..387

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5	CAC GAT GGG AAA GGA CAC CCT CTC CAC AGA AGA GAA AAG CGG CAA	48
	His Asp Gly Lys Gly His Pro Leu His Arg Arg Glu Lys Arg Gln	
	-10 -5 1	
96	A AAA CAC AAA CAG CGG AAA CGC CTC AAG TCC AGC TGT AAG AGA CAC	
	la Lys His Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His	
	5 10 15	
144	ECT TTA TAT GTG GAC TTC AGT GAT GTG GGG TGG AAT GAC TGG ATC GTT	
	Pro-Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val	
	20 25 30	
192	3CA CCG CCG GGG TAT CAT GCC TTT TAC TGC CAT GGG GAG TGC CCT TTT	
	Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly Glu Cys Pro Phe	
	35 40 45	
240	CCC CTG GCC GAT CAC CTT AAC TCC ACG AAT CAT GCC ATT CTC CAA ACT	
	Pro-Leu Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Leu Gln Thr	
	50 55 60 65	
288	CTG GTC AAC TCA GTT AAC TCT AAG ATT CCC AAG GCA TGC TGT GTC CCA	
	Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro	
	70 75 80	
336	ACA GAG CTC AGC GCC ATC TCC ATG CTG TAC CTT GAT GAG AAT GAG AAG	
	Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys	
	85 90 95	
384	GTG GTA TTA AAG AAC TAT CAG GAC ATG GTT GTC GAG GGT TGT GGG TGT	
	Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu Gly Cys Gly Cys	
	100 105 110	
437	CGT TAGCACAGCA AAATAAAATA TAAATATATA TATATATATA TTAGAAAAAC	
	Arg	
115		
497	AGCAAAAAAA TCAAGTTGAC ACTTTAATAT TTCCAATGA AGACTTTATT TATGGAATGG	
557	AATGGAGAAA AAGAAAAACA CAGCTATTTT GAAACTATA TTTATATCTA CCGAAAAGAA	
592	GTTGGGAAAA CAAATATTTT AATCAGAGAA TTATT	

INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 129 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gly His Asp Gly Lys Gly His Pro Leu His Arg Arg Glu Lys Arg Gln  
-15                      -10                      -5                      1  
Ala Lys His Lys Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His  
                    5                      10                      15  
Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val  
                    20                      25                      30  
Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly Glu Cys Pro Phe  
                    35                      40                      45  
Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Leu Gln Thr  
50                      55                      60                      65  
Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro  
                    70                      75                      80  
Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys  
                    85                      90                      95  
Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu Gly Cys Gly Cys  
100                      105                      110

Arg

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1607 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (G) CELL TYPE: Osteosarcoma Cell Line
- (H) CELL LINE: U-20S

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: U20S cDNA in Lambda GT10

(B) CLONE: L. Gaa U20S-39

(viii) POSITION IN GENOME:

(C) UNITS: bp

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 356..1546

(ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION: 1202..1543

(ix) FEATURE:

(A) NAME/KEY: mRNA

(B) LOCATION: 14..1607

(ix) FEATURE:

(A) NAME/KEY: sig\_peptide

(B) LOCATION: 356..424

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT CTTGAACTTG CAGGGAGAAT 60  
AACTTGCGCA CCCCACTTTG CGCCGGTGCC TTTGCCCCAG CGGAGCCTGC TTCGCCATCT 120  
CCGAGCCCCA CCGCCCCTCC ACTCCTCGGC CTTGCCCCGAC ACTGAGACGC TGTTCACAGC 180  
GTGAAAAGAG AGACTGCGCG GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG 240  
GACATTCGGT CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCATGTGGA CGCTCTTTCA 300  
ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC TCCTAAAGGT CGACC ATG 358  
Met  
-282  
GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC CTC 406  
Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val Leu  
-280 -275 -270  
CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG TTC 454  
Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe  
-265 -260 -255 -250  
GCG GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG GTC 502  
Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val  
-245 -240 -235  
CTG AGC GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG 550  
Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys Gln  
-230 -225 -220  
AGA CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA GAC 598  
Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu Asp  
-215 -210 -205  
CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC CAC 646

Leu	Tyr	Arg	Arg	His	Ser	Gly	Gln	Pro	Gly	Ser	Pro	Ala	Pro	Asp	His	
-200						-195					-190					
CGG	TTG	GAG	AGG	GCA	GCC	AGC	CGA	GCC	AAC	ACT	GTG	CGC	AGC	TTC	CAC	694
Arg	Leu	Glu	Arg	Ala	Ala	Ser	Arg	Ala	Asn	Thr	Val	Arg	Ser	Phe	His	
-185					-180				-175					-170		
CAT	GAA	GAA	TCT	TTG	GAA	GAA	CTA	CCA	GAA	ACG	AGT	GGG	AAA	ACA	ACC	742
His	Glu	Glu	Ser	Leu	Glu	Glu	Leu	Pro	Glu	Thr	Ser	Gly	Lys	Thr	Thr	
				-165					-160					-155		
CGG	AGA	TTC	TTC	TTT	AAT	TTA	AGT	TCT	ATC	CCC	ACG	GAG	GAG	TTT	ATC	790
Arg	Arg	Phe	Phe	Phe	Asn	Leu	Ser	Ser	Ile	Pro	Thr	Glu	Glu	Phe	Ile	
			-150					-145					-140			
ACC	TCA	GCA	GAG	CTT	CAG	GTT	TTC	CGA	GAA	CAG	ATG	CAA	GAT	GCT	TTA	838
Thr	Ser	Ala	Glu	Leu	Gln	Val	Phe	Arg	Glu	Gln	Met	Gln	Asp	Ala	Leu	
	-135					-130					-125					
GGA	AAC	AAT	AGC	AGT	TTC	CAT	CAC	CGA	ATT	AAT	ATT	TAT	GAA	ATC	ATA	886
Gly	Asn	Asn	Ser	Ser	Phe	His	His	Arg	Ile	Asn	Ile	Tyr	Glu	Ile	Ile	
-120					-115					-110						
AAA	CCT	GCA	ACA	GCC	AAC	TCG	AAA	TTC	CCC	GTG	ACC	AGA	CTT	TTG	GAC	934
Lys	Pro	Ala	Thr	Ala	Asn	Ser	Lys	Phe	Pro	Val	Thr	Arg	Leu	Leu	Asp	
-105				-100					-95					-90		
ACC	AGG	TTG	GTG	AAT	CAG	AAT	GCA	AGC	AGG	TGG	GAA	AGT	TTT	GAT	GTC	982
Thr	Arg	Leu	Val	Asn	Gln	Asn	Ala	Ser	Arg	Trp	Glu	Ser	Phe	Asp	Val	
			-85					-80					-75			
ACC	CCC	GCT	GTG	ATG	CGG	TGG	ACT	GCA	CAG	GGA	CAC	GCC	AAC	CAT	GGA	1030
Thr	Pro	Ala	Val	Met	Arg	Trp	Thr	Ala	Gln	Gly	His	Ala	Asn	His	Gly	
	-70			-65							-60					
ITC	GTG	GTG	GAA	GTG	GCC	CAC	TTG	GAG	GAG	AAA	CAA	GGT	GTC	TCC	AAG	1078
Phe	Val	Val	Glu	Val	Ala	His	Leu	Glu	Glu	Lys	Gln	Gly	Val	Ser	Lys	
	-55			-50							-45					
AGA	CAT	GTT	AGG	ATA	AGC	AGG	TCT	TTG	CAC	CAA	GAT	GAA	CAC	AGC	TGG	1126
Arg	His	Val	Arg	Ile	Ser	Arg	Ser	Leu	His	Gln	Asp	Glu	His	Ser	Trp	
-40				-35							-30					
ICA	CAG	ATA	AGG	CCA	TTG	CTA	GTA	ACT	TTT	GGC	CAT	GAT	GGA	AAA	GGG	1174
Ser	Gln	Ile	Arg	Pro	Leu	Leu	Val	Thr	Phe	Gly	His	Asp	Gly	Lys	Gly	
-25				-20				-15						-10		
CAT	CCT	CTC	CAC	AAA	AGA	GAA	AAA	CGT	CAA	GCC	AAA	CAC	AAA	CAG	CGG	1222
His	Pro	Leu	His	Lys	Arg	Glu	Lys	Arg	Gln	Ala	Lys	His	Lys	Gln	Arg	
			-5					1					5			
AAA	CGC	CTT	AAG	TCC	AGC	TGT	AAG	AGA	CAC	CCT	TTG	TAC	GTG	GAC	TTC	1270
Lys	Arg	Leu	Lys	Ser	Ser	Cys	Lys	Arg	His	Pro	Leu	Tyr	Val	Asp	Phe	
	10			15							20					
AGT	GAC	GTG	GGG	TGG	AAT	GAC	TGG	ATT	GTG	GCT	CCC	CCG	GGG	TAT	CAC	1318
Ser	Asp	Val	Gly	Trp	Asn	Asp	Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	His	
25				30							35					
GCC	TTT	TAC	TGC	CAC	GGA	GAA	TGC	CCT	TTT	CCT	CTG	GCT	GAT	CAT	CTG	1366

Ala Phe Tyr Cys His Gly [REDACTED] Cys Pro Phe Pro Leu Ala [REDACTED] Asp His Leu  
 40 45 50 55  
 AAC TCC ACT AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GTT AAC 1414  
 Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn  
 60 65 70  
 TCT AAG ATT CCT AAG GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC 1462  
 Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile  
 75 80 85  
 TCG ATG CTG TAC CTT GAC GAG AAT GAA AAG GTT GTA TTA AAG AAC TAT 1510  
 Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn Tyr  
 90 95 100  
 TAG GAC ATG GTT GTG GAG GGT TGT GGG TGT CGC TAGTACAGCA AAATTAAATA 1563  
 Gln Asp Met Val Val Glu Gly Cys Gly Cys Arg  
 105 110 115  
 ATAAATATA TATATATATA TATATTTTAG AAAAAAGAAA AAAA 1607

2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 396 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val  
 -282 -280 -275 -270  
 Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys  
 -265 -260 -255  
 He Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu  
 -250 -245 -240 -235  
 Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys  
 -230 -225 -220  
 Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu  
 -215 -210 -205  
 Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp  
 -200 -195 -190  
 His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe  
 -185 -180 -175  
 His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr  
 -170 -165 -160 -155  
 Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe  
 -150 -145 -140

Ile Thr Ser Ala Glu Leu Val Phe Arg Glu Gln Met Asn Asp Ala  
 -135 -130 -125  
 Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile  
 -120 -115 -110  
 Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu  
 -105 -100 -95  
 Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp  
 -90 -85 -80 -75  
 Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His  
 -70 -65 -60  
 Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser  
 -55 -50 -45  
 Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser  
 -40 -35 -30  
 Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys  
 -25 -20 -15  
 Gly His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His Lys Gln  
 -10 -5 1 5  
 Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp  
 10 15 17 20  
 Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr  
 25 30 35  
 His Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His  
 40 45 50  
 Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val  
 55 60 65 70  
 Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala  
 75 80 85  
 Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn  
 90 95 100  
 Tyr Gln Asp Met Val Val Glu Gly Cys Gly Cys Arg  
 105 110 114  
 382  
 396

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1954 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (G) CELL TYPE: Osteosarcoma Cell Line
- (H) CELL LINE: U-2OS

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: U2OS cDNA in Lambda gt10
- (B) CLONE: Lambda U2OS-3

(viii) POSITION IN GENOME:

- (C) UNITS: bp

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 403..1629

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 1279..1626

(ix) FEATURE:

- (A) NAME/KEY: mRNA
- (B) LOCATION: 9..1934

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTCTAGAGGG CAGAGGAGGA GGGAGGGAGG GAAGGAGCGC GGAGCCCGGC CCGGAAGCTA	60
GGTGAGTGTG GCATCCGAGC TGAGGGACGC GAGCCTGAGA CGCCGCTGCT GCTCCGGCTG	120
AGTATCTAGC TTGTCTCCCC GATGGGATTC CCGTCCAAGC TATCTCGAGC CTGCAGCGCC	180
ACAGTCCCCG GCCCTCGCCC AGGTTCACTG CAACCGTTCA GAGGTCCCCA GGAGCTGCTG	240
CTGGCGAGCC CGCTACTGCA GGGACCTATG GAGCCATTCC GTAGTGCCAT CCCGAGCAAC	300
GCACTGCTGC AGCTTCCCTG AGCCTTTCCA GCAAGTTTGT TCAAGATTGG CTGTCAAGAA	360
TCATGGACTG TTATTATATG CCTTGTTTTC TGTCAAGACA CC ATG ATT CCT GGT	414
Met Ile Pro Gly	
-292 -290	
AAC CGA ATG CTG ATG GTC GTT TTA TTA TGC CAA GTC CTG CTA GGA GGC	462
Asn Arg Met Leu Met Val Val Leu Leu Cys Gln Val Leu Leu Gly Gly	
-285 -280 -275	
GCG AGC CAT GCT AGT TTG ATA CCT GAG ACG GGG AAG AAA AAA GTC GCC	510
Ala Ser His Ala Ser Leu Ile Pro Glu Thr Gly Lys Lys Lys Val Ala	
-270 -265 -260	
GAG ATT CAG GGC CAC GCG GGA GGA CGC CGC TCA GGG CAG AGC CAT GAG	558
Glu Ile Gln Gly His Ala Gly Gly Arg Arg Ser Gly Gln Ser His Glu	
-255 -250 -245	
CTC CTG CGG GAC TTC GAG GCG ACA CTT CTG CAG ATG TTT GGG CTG CGC	606
Leu Leu Arg Asp Phe Glu Ala Thr Leu Leu Gln Met Phe Gly Leu Arg	



-240

-211

-230

-225

CGC CGC CCG CAG CCT AGC AAG AGT GCC GTC ATT CCG GAC TAC ATG CGG 654  
Arg Arg Pro Gln Pro Ser Lys Ser Ala Val Ile Pro Asp Tyr Met Arg  
-220 -215 -210

GAT CTT TAC CGG CTT CAG TCT GGG GAG GAG GAG GAA GAG CAG ATC CAC 702  
Asp Leu Tyr Arg Leu Gln Ser Gly Glu Glu Glu Glu Glu Gln Ile His  
-205 -200 -195

AGC ACT GGT CTT GAG TAT CCT GAG CGC CCG GCC AGC CGG GCC AAC ACC 750  
Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala Ser Arg Ala Asn Thr  
-190 -185 -180

GTG AGG AGC TTC CAC CAC GAA GAA CAT CTG GAG AAC ATC CCA GGG ACC 798  
Val Arg Ser Phe His His Glu Glu His Leu Glu Asn Ile Pro Gly Thr  
-175 -170 -165

AGT GAA AAC TCT GCT TTT CGT TTC CTC TTT AAC CTC AGC AGC ATC CCT 846  
Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu Ser Ser Ile Pro  
-160 -155 -150 -145

JAG AAC GAG GTG ATC TCC TCT GCA GAG CTT CGG CTC TTC CGG GAG CAG 894  
Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu Phe Arg Glu Gln  
-140 -135 -130

GTG GAC CAG GGC CCT GAT TGG GAA AGG GGC TTC CAC CGT ATA AAC ATT 942  
Val Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His Arg Ile Asn Ile  
-125 -120 -115

TAT GAG GTT ATG AAG CCC CCA GCA GAA GTG GTG CCT GGG CAC CTC ATC 990  
Tyr Glu Val Met Lys Pro Pro Ala Glu Val Val Pro Gly His Leu Ile  
-110 -105 -100

ACA CGA CTA CTG GAC ACG AGA CTG GTC CAC CAC AAT GTG ACA CGG TGG 1038  
Thr Arg Leu Leu Asp Thr Arg Leu Val His His Asn Val Thr Arg Trp  
-95 -90 -85

GAA ACT TTT GAT GTG AGC CCT GCG GTC CTT CGC TGG ACC CGG GAG AAG 1086  
Glu Thr Phe Asp Val Ser Pro Ala Val Leu Arg Trp Thr Arg Glu Lys  
-80 -75 -70 -65

CAG CCA AAC TAT GGG CTA GCC ATT GAG GTG ACT CAC CTC CAT CAG ACT 1134  
Gln Pro Asn Tyr Gly Leu Ala Ile Glu Val Thr His Leu His Gln Thr  
-60 -55 -50

CGG ACC CAC CAG GGC CAG CAT GTC AGG ATT AGC CGA TCG TTA CCT CAA 1182  
Arg Thr His Gln Gly Gln His Val Arg Ile Ser Arg Ser Leu Pro Gln  
-45 -40 -35

GGG AGT GGG AAT TGG GCC CAG CTC CGG CCC CTC CTG GTC ACC TTT GGC 1230  
Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val Thr Phe Gly  
-30 -25 -20

CAT GAT GGC CGG GGC CAT GCC TTG ACC CGA CGC CGG AGG GCC AAG CGT 1278  
His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg Arg Ala Lys Arg  
-15 -10 -5

AGC CCT AAG CAT CAC TCA CAG CGG GCC AGG AAG AAG AAT AAG AAC TGC 1326  
Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys

[illegible]

(2) INFORMATION FOR SEQ ID NO:6:

(ii) MOLECULE TYPE: protein

Met Ile Pro Gly Asn Arg Met Leu Met Val Val Leu Leu Cys Gln Val  
-292        -290                      -285                      -280

Gln Ser His Glu Leu Leu Arg Asp Phe Glu Ala Thr Leu Leu Gln Met  
 -240 -235 -230  
 Phe Gly Leu Arg Arg Arg Pro Gln Pro Ser Lys Ser Ala Val Ile Pro  
 -225 -220 -215  
 Asp Tyr Met Arg Asp Leu Tyr Arg Leu Gln Ser Gly Glu Glu Glu  
 -210 -205 -200  
 Glu Gln Ile His Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala Ser  
 -195 -190 -185  
 Arg Ala Asn Thr Val Arg Ser Phe His His Glu Glu His Leu Glu Asn  
 -180 -175 -170 -165  
 Ile Pro Gly Thr Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu  
 -160 -155 -150  
 Ser Ser Ile Pro Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu  
 -145 -140 -135  
 Phe Arg Glu Gln Val Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His  
 -130 -125 -120  
 Arg Ile Asn Ile Tyr Glu Val Met Lys Pro Pro Ala Glu Val Val Pro  
 -115 -110 -105  
 Gly His Leu Ile Thr Arg Leu Leu Asp Thr Arg Leu Val His His Asn  
 -100 -95 -90 -85  
 Val Thr Arg Trp Glu Thr Phe Asp Val Ser Pro Ala Val Leu Arg Trp  
 -80 -75 -70  
 Thr Arg Glu Lys Gln Pro Asn Tyr Gly Leu Ala Ile Glu Val Thr His  
 -65 -60 -55  
 Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser Arg  
 -50 -45 -40  
 Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu  
 -35 -30 -25  
 Val Thr Phe Gly His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg  
 -20 -15 -10 -5  
 Arg Ala Lys Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys  
 1 5 10  
 Asn Lys Asn Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val  
 15 20 25  
 Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr  
 30 35 40  
 Cys His Gly Asp Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr  
 45 50 55 60  
 Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile  
 65 70 75

